

## PICTURE RECOGNITION DEVICE

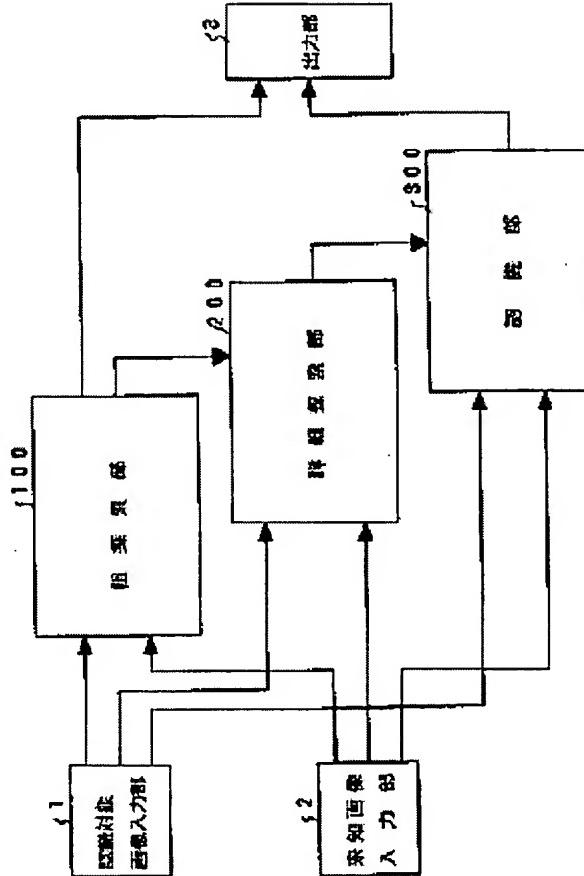
**Patent number:** JP5174149  
**Publication date:** 1993-07-13  
**Inventor:** KOSUGI MAKOTO  
**Applicant:** NIPPON TELEGRAPH & TELEPHONE  
**Classification:**  
 - international: G06T7/00; G06T7/00; (IPC1-7): G06F15/70  
 - european:  
**Application number:** JP19910344489 19911226  
**Priority number(s):** JP19910344489 19911226

[Report a data error here](#)

### Abstract of JP5174149

**PURPOSE:** To enable the recognition after an object picture is easily found and the position and size are obtained by making a recognition object picture and an unknown picture in a mosaic state and searching/recognizing them.

**CONSTITUTION:** A recognition object picture input part 1 fetches an object to be a recognition object, for instance, a typical human face picture if the object is human by using a TV camera or a scanner, etc., and inputs it in a rough search part 100. The object picture to be a recognition object is made in a rough mosaic state and rough position and size are found by scanning an unknown picture by using this picture as the feature of the object picture at the time of a search. Next, the object picture to be the recognition object is made in a fine mosaic state and accurate position and size are found by scanning around the already obtained rough position by using this picture. Further, recognition object picture is made in a fine mosaic and the object picture is recognized by performing the matching of the mosaic of the area obtaining the position and size in the already obtained unknown picture and the data which the previous recognition object picture is made in a finer mosaic.



Data supplied from the **esp@cenet** database - Worldwide

(10)日本特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平5-174149

(13)公開日 平成5年(1993)7月13日

(51)Int.Cl.  
G06F 15/70

識別記号  
450  
序内整理番号  
9071-5L

F I

技術表示箇所

審査請求 未請求 請求項の数8(全 11 頁)

(21)出願番号 特願平3-344489

(71)出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(22)出願日 平成3年(1991)12月28日

(72)発明者 小林 信

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

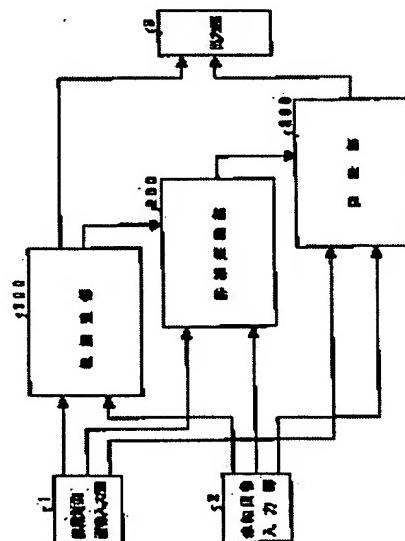
(74)代理人弁理士 稲田 実

(54)【発明の名前】 画像認識装置

(57)【要約】

【目的】 本発明は、制約条件を緩和し、目的とする画像を画面内のどこにあるかを正確に求め、この結果に基づき対象を認識できるようにすることを目的としている。

【構成】 認識対象となる物体画像を粗くモザイク化し、これを探索時の対象画像の特徴として用いて、未知画像を走査しおよその位置と大きさとを見つける。次に、認識対象となる物体画像を細かくモザイク化し、これを用いて、既に得られた大まかな位置付近を走査し正確な位置と大きさとを見つける。さらに、認識対象物体画像を細かくモザイク化し、未知画像のモザイクとのマッチングをとる。



【特許請求の範囲】

【請求項 1】 予め、認識対象となる自然画像を大きなブロックサイズで $M \times N$ に分割し、各ブロックの邊淡あるいはカラーの代表値を算出して辞書データとして記憶する第1の手段と、  
同じく、予め、当該認識対象の自然画像を小さなブロックサイズで $M' \times N'$ に分割し、各ブロックの邊淡あるいはカラーの代表値を算出して辞書データとして記憶する第2の手段と、  
同じく、予め当該認識対象の自然画像をさらに小さなブロックサイズで $M'' \times N''$ に分割し、各ブロックの邊淡あるいはカラーの代表値を算出して辞書データとして記憶する第3の手段と、  
未知の自然画像が入力として与えられたとき、入力画像を大きなブロックサイズで $P \times Q$ に分割して各ブロックの代表値を算出し、この中から任意の $M \times N$ の領域を取り出して第1の手段で得られた代表値群とのなす距離を算出し審査する第4の手段と、  
 $P \times Q$ の全領域にわたって第4の手段を繰り返し適用し、得られた距離の中から最小値を求め、この最小値が予め与えられた閾値より小さいとき、入力画像中の最小値を導いた位置に認識対象画像に相当する画像があると判定し、一方、最小値が閾値より大きいとき、入力画像内に認識対象画像に相当する画像が存在しないと判断する第5の手段と、  
第5の手段で、或る位置に認識対象有りと判定されたとき、入力画像中の該当位置の近傍の $M' \times N'$ 領域を小さなブロックで $P' \times Q'$ に分割し各ブロックの代表値を算出し、この中から任意の $M'' \times N''$ の領域を取り出して第2の手段で得られた代表値群とのなす距離を算出しこれを審査する第6の手段と、  
 $P' \times Q'$ の全領域にわたって第6の手段を繰り返し適用し、得られた距離の中から最小値を求め、この時の $M'' \times N''$ 画像の位置と大きさを認識対象画像の位置と大きさとする第7の手段と、  
第7の手段で得られた入力画像中の該当位置と大きさの領域を、より小さなブロックサイズで $M'' \times N''$ に分割して各ブロックの代表値を算出し、第3の手段で得られた代表値群とのなす距離を算出し、これが或る閾値以下ならば認識対象画像であると特定する第8の手段を有することを特徴とする画像認識装置。  
【請求項 2】 上記第1の手段において、予め対象となる画像を唯一でなく、類似する複数の画像群とし、これらの画像ごとに $M \times N$ に分割して各ブロックの代表値を算出し、複数の画像間の対応するブロックごとに平均値を算出し、これを辞書として記憶することを特徴とする請求項1記載の画像認識装置。  
【請求項 3】 上記第2の手段において、予め対象となる画像を唯一でなく、類似する複数の画像群とし、これらの画像ごとに $M' \times N'$ に分割して各ブロックの代表

値を算出し、複数の画像間の対応するブロックごとに平均値を算出し、これを辞書として記憶することを特徴とする請求項1記載の画像認識装置。

【請求項 4】 上記第3の手段において、予め対象となる画像を唯一でなく、複数の画像群とし、これらの画像ごとに $M'' \times N''$ に分割して各ブロックの代表値を算出し、これを辞書として記憶することを特徴とする請求項1記載の画像認識装置。

【請求項 5】 上記第2の手段において、予め、認識対象画像の部分画像を小さなブロックサイズで $M' \times N'$ に分割し、各ブロックの代表値を算出して辞書データとして記憶し、上記第5の手段および第7の手段において、上記第5の手段で得られた $M \times N$ 領域の部分領域を $P' \times Q'$ に分割し、この中の $M'' \times N''$ ごとに当該辞書データとの距離を算出することを特徴とする請求項1記載の画像認識装置。

【請求項 6】 上記第3の手段において、予め、認識対象画像の部分画像を、より小さなブロックサイズで $M'' \times N''$ に分割し、各ブロックの代表値を算出して辞書データとして記憶し、上記第8の手段において、上記 $M'' \times N''$ 領域を、より小さなブロックサイズで $M''' \times N'''$ に分割することを特徴とする請求項1記載の画像認識装置。

【請求項 7】 上記第4の手段において、分割に用いるブロックサイズを可変とすることを特徴とする請求項1記載の画像認識装置。

【請求項 8】 上記第6の手段において、分割に用いるブロックサイズを可変とすることを特徴とする請求項1記載の画像認識装置。

【発明の詳細な説明】

【00101】

【産業上の利用分野】 本発明は、入力された自然画像の中から目的とする画像を精度よく探し、対象を特定するようにした画像認識装置に関する。

【00102】

【従来の技術】 代表的な自然画像は邊淡あるいはカラー画像であるが、自然画像の中から目的とする画像を検出し認識するため、従来は、画像の形状に注目し、形状情報を特徴として入力画像内から特徴の一一致する候補を探出し、認識するものであった。しかし、自然画像の中から正しい形状を抽出することは至難であり、従来は、対象物体と背景の切り分けのため、背景を事前学習したりすること、あるいは背景は一様なものに制限すること、複数を抽出し見くするため人工的な剛物体などに対象を制限すること、など種々の制約条件を課していた。

【00103】

【発明が解決しようとする課題】 目的の画像の有無の検知できこのような問題が伴っており、まして自然画像において対象物の正確な位置や大きさをつかむことは不可能に近かった。

【0004】即ち、任意の自然画像の中から対象物を正確に抽出することは至難のことであり、このため、実用上、割合が多く有用性に難があった。本発明は、上記のさまざまな制約条件を緩和し、目的とする画像を画面内のどこにあるか正確に求め、この結果に基づき対象を認識できるようにすることを目的としている。

【0005】

【課題を解決するための手段】本発明においては、対象の特徴を形状ではなく粗い解像度の追徴あるいはカラー情報を用いる。即ち、リンクーンのモザイク画像からリンクーンが識別できるように、解像度を大きく落とした画像でも、対象の特徴を表現できることを根拠としている。

【0006】具体的には、認識対象となる物体画像を粗くモザイク化し、これを探索時の対象画像の特徴として用いて、未知画像を走査しおよその位置と大きさを見つける。次に、認識対象となる物体画像を細かくモザイク化し、これを用いて、既に得られた大まかな位置附近を走査し正確な位置と大きさを見つける。さらに、認識対象物体画像を細かくモザイク化し、すでに得られた未知画像中の位置と大きさを得ている領域のモザイクと上記先の認識対象物体画像を細かくモザイクしたデータのマッチングにより対象画像を認識する。

【0007】

【作用】このように画像を数段階でモザイク化することにより、任意の画面中から対象を正確に探し出し、さらに認識することが容易に可能となる。

【0008】

【実施例】以下、本発明の一実施例について図面により説明する。図1は本発明における画像認識を行う一実施例を示す図であり、図2は図1に示す粗探索部の構成、図3は図1に示す詳細探索部の構成、図4は図1に示す認識部の構成を示す。また図5と図6とは一緒にになって1つの図を表しており探索対象画像あるいは未知画像をモザイク化し、未知画像の中から目的的対象を探索・認識する過程を説明する図である。

【0009】図1において1は認識対象画像入力部であり、TVカメラあるいはスキャナなどを用い、認識対象となる物体、例えば人物であれば典型的な人物の顔画像を取り込み、粗探索部100へ、具体的には、図2の粗探索対象画像バッファ101へ入力する。したがって、当該バッファ101の内容は追従あるいはカラーで表される画素の集合である。

【0010】この認識対象画像を行列 $F = [f]$ で表し、図5、図6のように、これを $W_e$ 画素× $W_e$ 画素のサイズのブロックで $M \times N$ に分割し、このモザイク画像 $F_C$ の各ブロックをブロック内の代表値、例えば追従ならばブロック内平均値、カラーならばブロック内で最大の輝度を有する色です。

【0011】粗探索対象画像モザイク化部102は、前

記バッファ101の画像をモザイク化し、粗探索対象画像辞書103に蓄積する。なお、必要ならば、複数の探索対象、この例では複数の顔画像の平均値をとり、代表的な顔のモザイクデータとするか、複数の粗探索対象画像を用意してもよい。

【0012】一方、未知画像 $I$ は図1の未知画像入力部2を介して粗探索部100の中の、図2に示す未知画像バッファ104に取り込まれる。そこで未知画像 $I$ に対して、やはり図5のように、未知画像モザイク化部105で $W_e$ 画素× $W_e$ 画素のブロックサイズで $P \times Q$ に分割し、ブロックごとに代表値を算出しモザイク画像 $U_C$ を得る。なお、未知画像入力部2は前記の認識対象画像入力部1と同じでもよい。

【0013】次に、図5のように、未知画像のモザイクを走査して粗探索対象画像モザイクヒー致する場所を探索する。即ち、次の原理に基づく。未知画像の $P \times Q$ のモザイクデータのうち、任意の $M \times N$ のモザイクデータを $U_1$ （斜線が存在する領域で示す）としたとき、 $U_1$ と粗探索対象モザイクデータ $F_C$ との距離 $D_1$ が最小値となるときの $U_1$ の位置を粗探索結果とする。この距離 $D_1$ として、種々の出し方があるが代表的なものとしてユーリッド距離がある。

【0014】具体的には次のとおりである。図2において、106は候補選択部であり、未知画像モザイク化部105の中から $M \times N$ ブロック分の枠に入る候補データ $U_i$  ( $i = 1, 2, \dots$ ) を選択する。一般的には、はじめに未知画像の左上の $M \times N$ ブロック分が $U_1$ として選択される。このモザイクデータ $U_1$ との粗探索対象画像辞書103からの粗探索対象モザイクデータ $F_C$ との距離算出部107に入力され、上記の距離が算出される。この結果は、このときの未知画像モザイクデータ $U_C$ における $M \times N$ の位置データ、即ち $U_1$ の位置データ、ならびにモザイクのブロックサイズ $W_e$ とともに位置・サイズ・距離蓄積部108に蓄えられる。

【0015】さらに、候補選択部106は未知画像モザイク化部105から次の候補 $U_2$ を選択し、距離算出部107はこの時の距離を算出して、候補 $U_2$ の位置・ブロックサイズとともに位置・サイズ・距離蓄積部108に蓄える。同様に、未知画像モザイクの全領域に対して $M \times N$ の候補 $U_1$ が順次選択され、選択された位置・ブロックサイズと算出された距離はすべて位置・サイズ・距離蓄積部108に蓄えられる。

【0016】ところで、以上では未知画像のモザイク化において、探索対象画像と同じブロックサイズ、 $W_e$ 画素× $W_e$ 画素を用いた。探索対象画像と未知画像内の探索対象画像が同じサイズであればこのままでよいが、一般的には、未知画像内における探索対象画像のサイズもまた未知である。このため、未知画像モザイク化部105は、ブロックサイズの画素数を、例えば、 $W_e$ の $1/5 \sim 1/10 \sim 5$ などの $W_e'$ に変化させる。そこで、一つ

一つのブロックサイズ  $W_a'$  ことに上記の手順を繰り返して、得られた距離とそのときの位置ならびにブロックサイズ  $W_a'$  を位置・サイズ・距離寄積部 108 に蓄える。

【0017】また、複数の探索対象画像を辞書においていた場合も、同様にして、一つ一つの探索対象画像ごとに上記の手順を繰り返し、得られた距離とそのときの位置およびブロックサイズを位置・サイズ・距離寄積部 108 に蓄える。

【0018】こうして、全部の場合の距離が位置・サイズ・距離寄積部 108 に蓄えられると、最小距離検出部 109 と、位置・サイズ・距離寄積部 108 の中の距離データのうち、最小値を取りものを検出す。この値は認識対象有無判定部 110 に送られ、この値がある閾値以下の時は未知画像内に対象画像があると判断し、その位置およびブロックサイズを位置・サイズ・距離寄積部 108 から取り出して結果を粗探索結果出力部 111 に出力する。粗探索結果出力部 111 は、ブロックサイズ  $W_a$  から未知画像内の探索対象画像の大きさを  $[M \times W_a] \times [N \times W_a]$  より算出する。

【0019】一方、最小距離値がある閾値を超えたときは未知画像内に対象画像が無いと判断し、「無い」ことを粗探索結果出力部 111 に出力する。粗探索結果出力部 111 はこれらの結果を直ちに詳細探索部 200 に通知する。

【0020】こうして粗探索により、認識対象のおおよその位置と大きさがわかると、モザイクのブロックを小さくして、詳細にその位置を探索する。なお、詳細探索では粗探索に比べ対象画像の領域を一部分に絞ってもいいが、ここでは、同一領域を用いる場合を述べる。

【0021】まず、予め、図 6 のように、認識対象を  $W_b$  画素  $\times$   $W_b$  画素 ( $W_b < W_a$ ) のサイズのブロックで  $M' \times N'$  に分割し、各ブロックをブロック内の代表値です。

【0022】具体的には、詳細探索対象画像バッファ 201 は、認識対象画像入力部 1 から認識対象画像データを取り込み、これを詳細探索対象画像モザイク化部 202 に送る。詳細探索対象画像モザイク化部 202 は、この画像をモザイク化 (FD) し、詳細探索対象画像辞書 203 に蓄積する。

【0023】なお、必要ならば、複数の詳細探索対象画像の平均値をとるか、複数の詳細探索対象画像を用意すればよいのは、粗探索の場合と同じである。一方、上記の粗探索結果が粗探索結果通知部 220 に通知されると、未知画像バッファ 204 はこれを知って、図 1 の未知画像入力部 2 から、未知画像 I 中の認識対象画像の該当部分を取り込む。そこで、未知画像モザイク化部 205 は、これに対して、やはり図 6 のように、 $W_b$  画素  $\times$   $W_b$  画素のブロックサイズで  $P' \times Q'$  に分割し、ブロックごとに代表値を算出しモザイク UDO を得る。

【0024】そこで、粗探索の場合と同様、図 6 のように、未知画像の該当部分画像のモザイクを走査して詳細探索対象モザイク FDO と一致する場所を探索する。すなわち、部分画像の  $P' \times Q'$  のモザイク UDO のうち任意の  $M' \times N'$  のモザイクを  $U'$  (斜線が存在する領域で示す) としたとき、 $U'$  と探索対象のモザイク FDO との距離  $D'$  を最小となるときの  $U'$  の位置を詳細探索結果とする。

【0025】具体的には、図 3において、候補選択部 206 は、未知画像モザイク化部 203 の中から  $M' \times N'$  分の内に入る候補  $U'$  を取り出す。距離算出部 207 は、この  $U'$  と詳細探索対象画像辞書 203 からの詳細探索対象モザイク FDO の距離  $D'$  を算出し、このときの  $U'$  の位置データ、ならびにモザイクのブロックサイズ  $W_b$  とともに位置・サイズ・距離寄積部 208 に送る。

【0026】ここで、必要に応じ、ブロックサイズ  $W_b$  を変化させ ( $W_b'$ ) たり、複数の詳細探索対象画像を用いたりすることも、粗探索の場合と同様である。こうして、最小距離検出部 209 で最小値を検出すると、この時の  $U'$  の位置データならびにブロックサイズ  $W_b$  が詳細探索結果出力部 210 に送られる。詳細探索結果出力部 210 はブロックサイズ  $W_b$  より探索対象画像の正確な大きさ、即ち、 $[M' \times W_b] \times [N' \times W_b]$  を算出し、位置データとともに認識部 300 に送出する。

【0027】こうして詳細探索により、認識対象の正確な位置と大きさがわかると、モザイクのブロックを小さくして、対象が何かを認識する。まず、予め、図 6 のように、認識対象を  $W_b$  画素  $\times$   $W_b$  画素 ( $W_b < W_b'$ ) のサイズのブロックで  $M'' \times N''$  に分割する。このモザイクの各ブロックを、付記の場合は同様にブロック内の代表値で表す。なお、認識対象の特徴が部分領域に集中しているれば、その部分を  $M'' \times N''$  に分割してもよい。すなわち、探索の過程では対象に普遍的な特徴が必要であるが、認識の過程では、個々を区別する特徴が必要であり、より細かなモザイクや部分領域が有効となる。

【0028】ここでは、この部分領域を用いる場合を述べる。具体的には、認識対象画像バッファ 301 は、認識対象画像入力部 1 から認識対象画像の部分領域のデータを取り込み、これを認識対象画像モザイク化部 302 に送る。認識対象画像モザイク化部 302 は、この画像をモザイク化 (FDD) し、認識対象画像辞書 303 に蓄積する。

【0029】なお、通常、多くの認識対象画像についてモザイク化し、認識対象画像辞書 303 に蓄積する。一方、上記の詳細探索結果が詳細探索部 200 より詳細探索結果通知部 320 に通知されると、未知画像バッファ 304 はこれを知って、図 1 の未知画像入力部 2 から、未知画像 I 中の認識対象画像の正確な位置と大きさの画像を取り込む。画像切り出し・モザイク化部 305

は、さらにその中の部分領域に対して、やはり図6のように $W_c$ 画素× $W_o$ 画素のブロックサイズで $M'' \times N''$ に分割し、ブロックごとに代表値を算出しモザイクデータUDDを得る。

【0030】そこで、このモザイクデータUDDと認識対象画像モザイクデータFDDの距離を算出し、これが或る閾値以下であれば、未知画像は認識対象画像であると判定する。あるいは、認識対象画像が複数あれば、FDDとの距離の最小となるものを認識対象として判定する。ここでは、一般的な、後者の場合を例に述べる。

【0031】具体的には、図4において、距離算出部306は、画像切り出し・モザイク化部303の出力である未知画像モザイクデータUDDと認識対象画像群書303からの認識対象画像モザイクデータFDDの距離を算出し、このときの距離と認識対象画像の符号を距離蓄積部307に送る。同時に、距離算出部306は、認識対象画像群書303から次々と認識対象画像モザイクデータを取り出しては、未知画像モザイクデータUDDとの距離を算出し、この値と認識対象画像の符号を距離蓄積部307に送る。

【0032】つぎに、最小距離検出部308は、距離蓄積部307の中の最小値を検出すると、この時の認識対象画像の符号を認識結果出力部309へ送出する。この結果はさらに出力部3へ送出される。

【0033】一方、粗探索部の結果、対象画像が「無し」の場合は、このことが出力部3に通知される。こうして、未知画像Uの中から、まず認識対象画像Fのおおよその位置がみつかり、これを手がかりに正確な位置と大きさが見つけられ、さらに対象が特定される。

#### 【0034】

【発明の効果】本発明によれば、認識対象画像および未知画像をモザイク化し探索・認識することによって、背景に対する一様性など特殊な条件を与えることなく、また、線分を用いた場合のノイズによるエラーを起すことなく、容易に対象画像を見いたしてその位置や大きさを得た上、認識することが可能となる。とくに、形状や色が似通った対象、例えば、人間の頭や姓などには効果的である。

#### 【図面の簡単な説明】

【図1】本発明の全体を説明するブロック図である。

【図2】粗探索部1-0の内容を示す図である。

【図3】詳細探索部2-0の内容を示す図である。

【図4】認識部3-0の内容を示す図である。

【図5】認識対象画像ならびに未知画像をモザイク化し、未知画像の中から認識対象を段階的に探索・認識する過程を説明する図である。

【図6】図5と一緒にになって1つの図を構成し、認識対象画像ならびに未知画像をモザイク化し、未知画像の中から認識対象を段階的に探索・認識する過程を説明する図である。

#### 【符号の説明】

- 1 認識対象画像入力部
- 2 未知画像入力部
- 3 出力部
- 1'00: 粗探索部
- 2'00: 詳細探索部
- 3'00: 認識部

【図1】

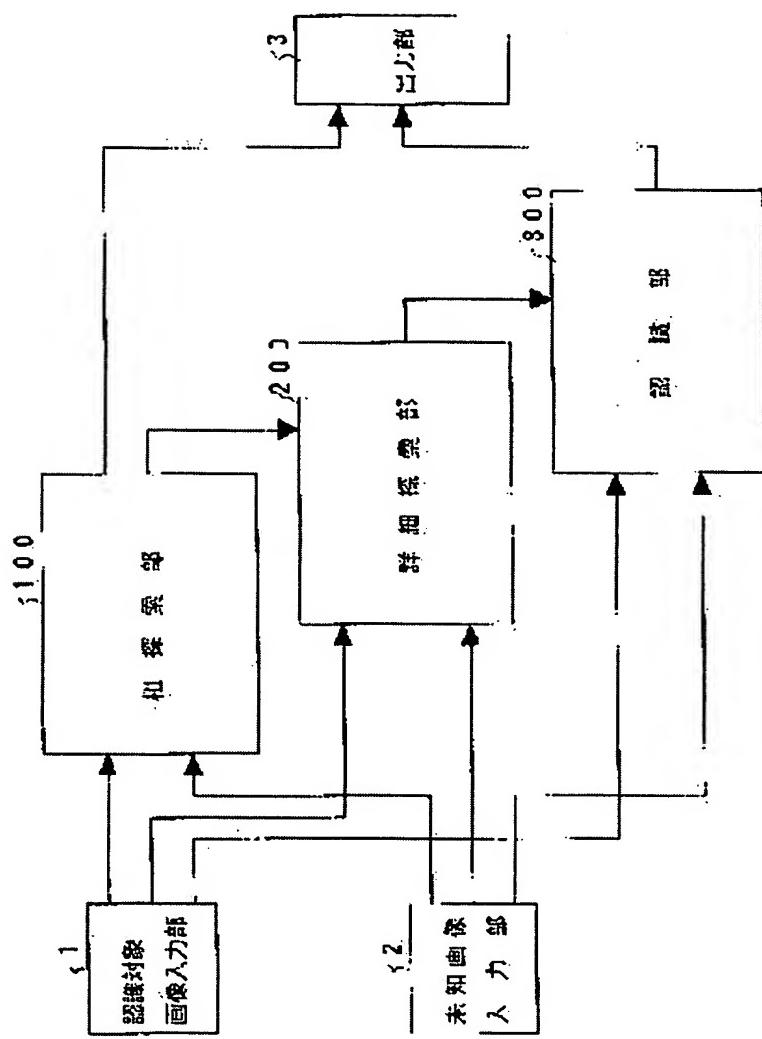
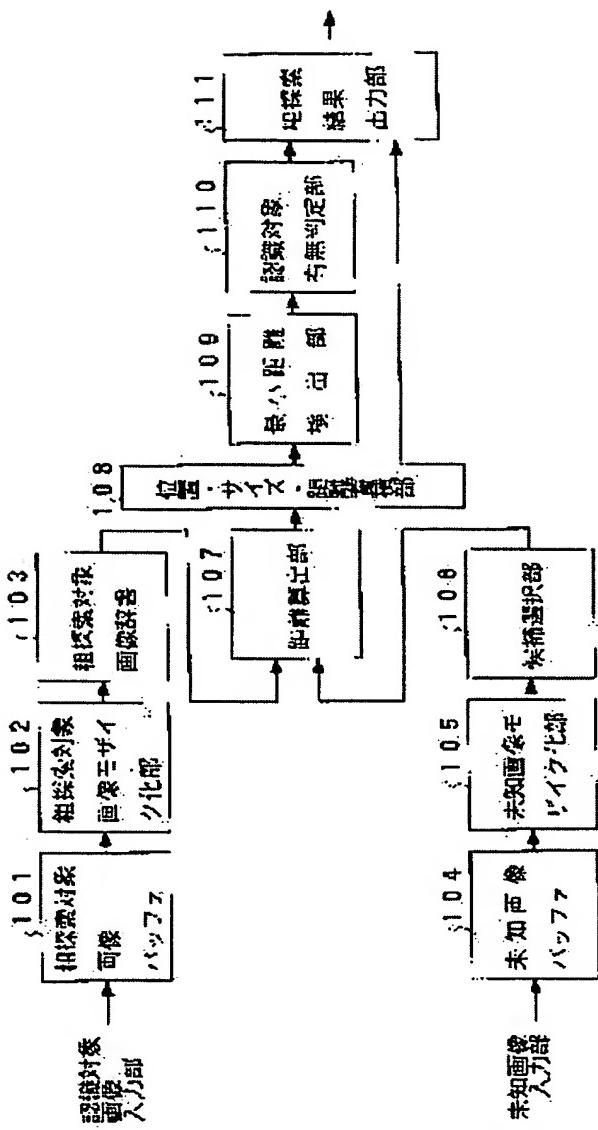
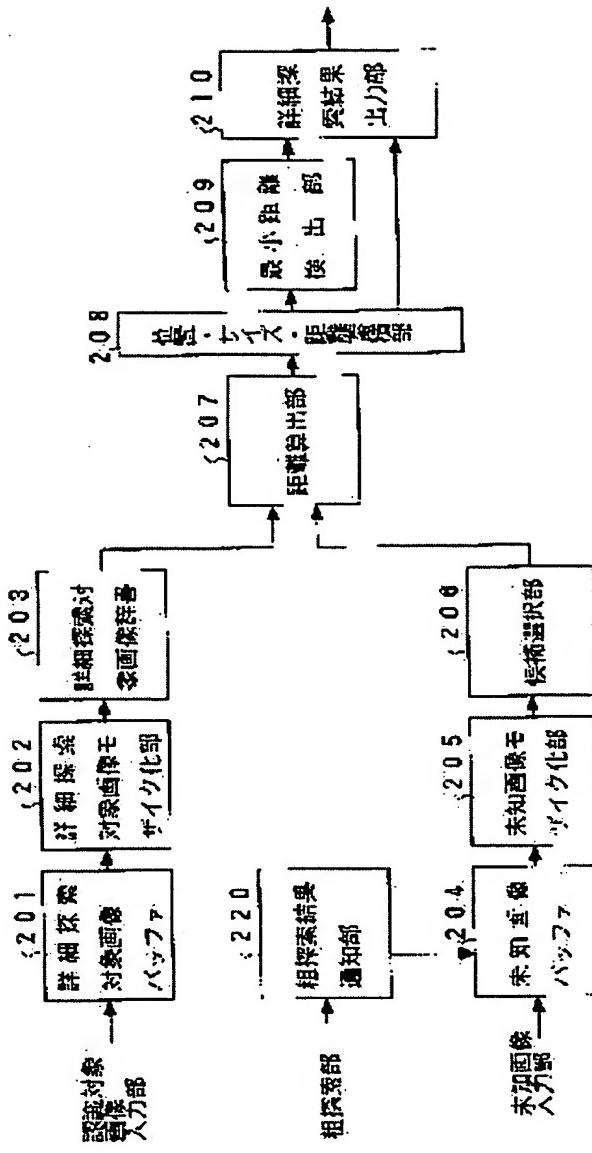


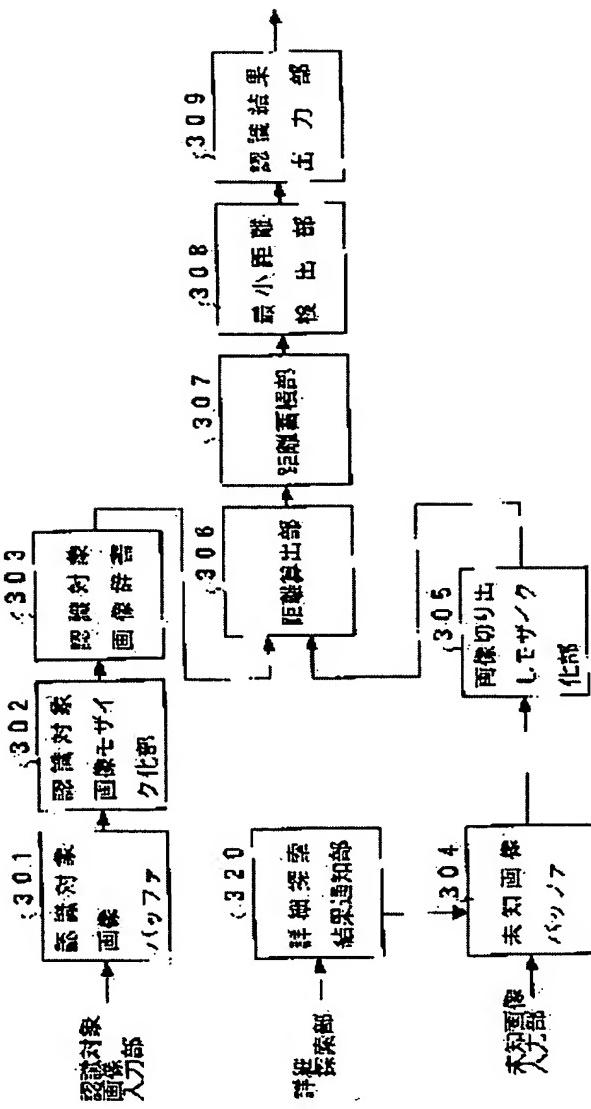
図2



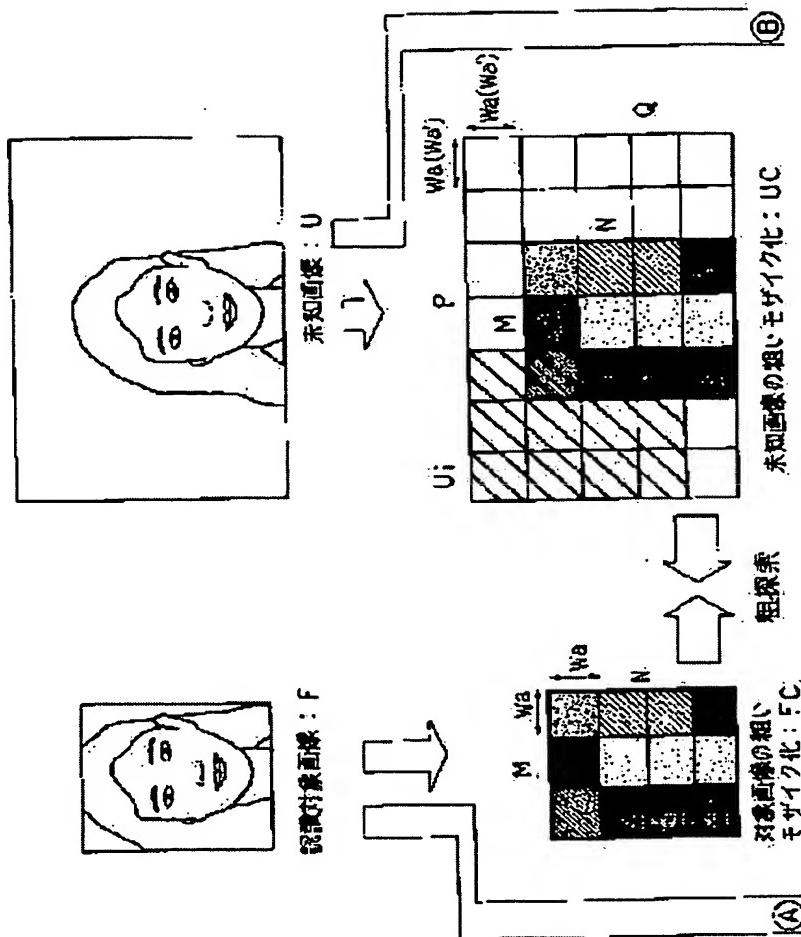
〔圖3〕



【図4】

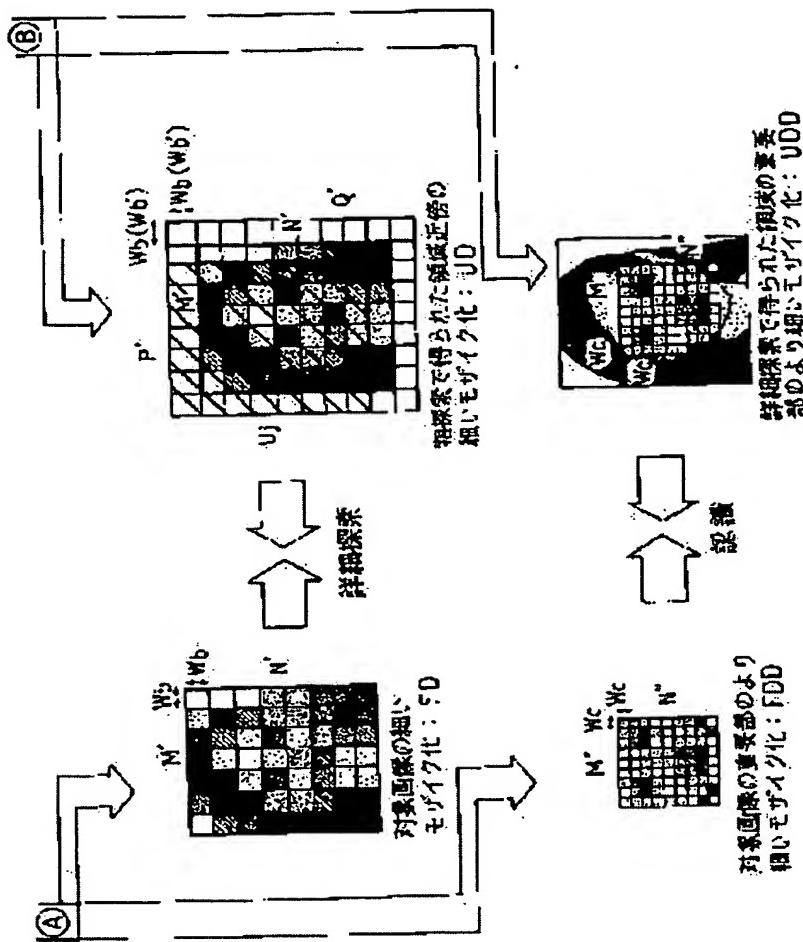


【図5】



BEST AVAILABLE COPY

【図6】



BEST AVAILABLE COPY